

#7

OIEP

RAW SEQUENCE LISTING

DATE: 07/18/2001

PATENT APPLICATION: US/09/745,763

TIME: 12:08:34

Input Set : A:\Cpg.pto

Output Set: N:\CRF3\07182001\I745763.raw

SEQUENCE LISTING

```

4 (1) GENERAL INFORMATION:
6     (i) APPLICANT: Jacobs, Kenneth
7         McCoy, John M.
8         LaVallie, Edward R.
9         Collins-Racie, Lisa A.
10        Evans, Cheryl
11        Merberg, David
12        Treacy, Maurice
13        Spaulding, Vikki
15     (ii) TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
16                               ENCODING THEM
18     (iii) NUMBER OF SEQUENCES: 219
20     (iv) CORRESPONDENCE ADDRESS:
21         (A) ADDRESSEE: Genetics Institute, Inc.
22         (B) STREET: 87 CambridgePark Drive
23         (C) CITY: Cambridge
24         (D) STATE: MA
25         (E) COUNTRY: U.S.A.
26         (F) ZIP: 02140
28     (v) COMPUTER READABLE FORM:
29         (A) MEDIUM TYPE: Floppy disk
30         (B) COMPUTER: IBM PC compatible
31         (C) OPERATING SYSTEM: PC-DOS/MS-DOS
32         (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
34     (vi) CURRENT APPLICATION DATA:
C--> 35         (A) APPLICATION NUMBER: US/09/745,763
C--> 36         (B) FILING DATE: 18-Jun-2000
37         (C) CLASSIFICATION:
39     (viii) ATTORNEY/AGENT INFORMATION:
40         (A) NAME: Sprunger, Suzanne A.
41         (B) REGISTRATION NUMBER: 41,323
43     (ix) TELECOMMUNICATION INFORMATION:
44         (A) TELEPHONE: (617) 498-8284
45         (B) TELEFAX: (617) 876-5851
48 (2) INFORMATION FOR SEQ ID NO: 1:
50     (i) SEQUENCE CHARACTERISTICS:
51         (A) LENGTH: 1800 base pairs
52         (B) TYPE: nucleic acid
53         (C) STRANDEDNESS: double
54         (D) TOPOLOGY: linear
55     (ii) MOLECULE TYPE: cDNA
60     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
62 TTTTTTTTTT TACAGACTTC ACAGAGAATG CAGTTGTCTT GACTTCAGGT CTGTCTGTTC      60
64 TGTTGGCAAG TAAATGCAGT ACTGTTCTGA TCCCGCTGCT ATTAGAATGC ATTGTGAAAC      120
66 GACTGGAGTA TGATTAAAG TTGTGTTCCC CAATGCTTGG AGTAGTGATT GTTGAAGGAA      180
68 AAAATCCAGC TGAGTGATAA AGGCTGAGTG TTGAGGAAAT TTCTGCAGTT TTAAGCAGTC      240

```

RAW SEQUENCE LISTING

DATE: 07/18/2001

PATENT APPLICATION: US/09/745,763

TIME: 12:08:34

Input Set : A:\Cpg.pto

Output Set: N:\CRF3\07182001\I745763.raw

```

70 GTATTTGTGA TTGAAGCTGA GTACATTTTG CTGGTGTATT TTTAGGTAAA ATGCTTTTTG 300
72 TTCATTTCTG GTGGTGGGAG GGGACTGAAG CCTTTAGTCT TTTCCAGATG CAACCTTAAA 360
74 ATCAGTGACA AGAAACATTC CAAACAAGCA ACAGTCTTCA AGAAATTAAA CTGGCAAGTG 420
76 GAAATGTTTA AACAGTTCAG TGATCTTTAG TGCATTGTTT ATGTGTGGGT TTCTCTCTCC 480
78 CCTCCCTTGG TCTTAATTCT TACATGCAGG AACACTCAGC AGACACACGT ATGCGAAGGG 540
80 CCAGAGAAGC CAGACCCAGT AAGAAAAAAT AGCCTATTTA CTTTAAATAA ACCAAACATT 600
82 CCATTTTAAA TGTGGGGATT GGGAACCACT AGTTCTTTCA GATGGTATTC TTCAGACTAT 660
84 AGAAGGAGCT TCCAGTTGAA TTCACCAGTG GACAAAATGA GGAAAACAGG TGAACAAGCT 720
86 TTTTCTGTAT TTACATACAA AGTCAGATCA GTTATGGGAC AATAGTATTG AATAGATTTT 780
88 AGCTTTATGC TGGAGTAACT GGCATGTGAG CAAACTGTGT TGGCGTGGGG GTGGAGGGGT 840
90 GAGGTGGGCG CTAAGCTTTT TTTAAGATTT TTCAGGTACC CTTCACTAAA GGCACCGAAG 900
92 GCTTAAAGTA GGACAACCAT GGAGCTTCCT GTGGCAGGAG AGACAACAAA GCGCTATTAT 960
94 CCTAAGGTCA AGAGAAGTGT CAGCCTCACC TGATTTTTAT TAGTAATGAG GACTTGCCTC 1020
96 AACTCCCTCT TTCTGGAGTG AAGCATCCGA AGGAATGCTT GAAGTACCCC TGGGCTTCTC 1080
98 TTAACATTTA AGCAAGCTGT TTTTATAGCA GCTCTTAATA ATAAAGCCCA AATCTCAAGC 1140
100 GGTGCTTGAA GGGGAGGGAA AGGGGGAAAG CGGGCAACCA CTTTTCCCTA GCTTTTCCAG 1200
102 AAGCCTGTTA AAAGCAAGGT CTCCCCACAA GCAACTTCTC TGCCACATCG CCACCCCGTG 1260
104 CCTTTTGATC TAGCACAGAC CCTTCACCCC TCACCTCGAT GCAGCCAGTA GCTTGGATCC 1320
106 TTGTGGGCAT GATCCATAAT CGGTTTCAAG GTAACGATGG TGTCGAGKTC TTTGGTGGGT 1380
108 TGAAGTATGT TAGAAAGGC CATTAAATTTG CCTGCAAATT GTTAAACAGAA GGGTATTAAA 1440
110 ACCACAGCTA AGTAGTCTA TTATAATACT TATCCAGTGA CTAAAACCAA CTTAAACCAG 1500
112 TAAGTGGAGA AATAACATGT TCAAGAACTG TAATGCTGGG TGGGAACATG TAACTTGTAG 1560
114 ACTGGAGAAG ATAGGCATTT GAGTGGCTGA GAGGGCTTTT GGGTGGGAAT GCAAAAATTC 1620
116 TCTGCTAAGA CTTTTTCAGG TGAACATAAC AGACTTGGCC AAGCTAGCAT CTTAGCGGAA 1680
118 GCTGATCTCC AATGCTCTTC AGTAGGGTCA TGAAGGTTTT TCTTTTCCTG AGAAAACAAC 1740
120 ACGTATTGTT TTCTCAGGTT TTGCTTTTTG GCCTTTTTCT AGCTTAAAAA AAAAAAAAAA 1800

```

124 (2) INFORMATION FOR SEQ ID NO: 2:

126 (i) SEQUENCE CHARACTERISTICS:

127 (A) LENGTH: 48 amino acids

128 (B) TYPE: amino acid

129 (C) STRANDEDNESS:

130 (D) TOPOLOGY: linear

132 (ii) MOLECULE TYPE: protein

135 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

137 Val Trp Val Ser Leu Ser Pro Pro Leu Val Leu Ile Leu Thr Cys Arg

138 1 5 10 15

140 Asn Thr Gln Gln Thr HisVal Cys Glu Gly Pro Glu Lys Pro Asp Pro

141 20 25 30

143 Val Arg Lys Asn Ser Leu Phe Thr Leu Asn Lys Pro Asn Ile Pro Phe

144 35 40 45

148 (2) INFORMATION FOR SEQ ID NO: 3:

150 (i) SEQUENCE CHARACTERISTICS:

151 (A) LENGTH: 1063 base pairs

152 (B) TYPE: nucleic acid

153 (C) STRANDEDNESS: double

154 (D) TOPOLOGY: linear

156 (ii) MOLECULE TYPE: cDNA

159 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

161 AAAGTTCCAT CTCTAGAACT GATTTTTATC CGTTCTGTTT TTCAGGTCTT ATCTGTGTTA 60

RAW SEQUENCE LISTING

DATE: 07/18/2001

PATENT APPLICATION: US/09/745,763

TIME: 12:08:34

Input Set : A:\Cpg.pto

Output Set: N:\CRF3\07182001\I745763.raw

```

163 GTTGTGTGTT ACTATCAGGA GGCCCCCTTT GGACCCAGTG GATACAGATT ACGACTCTTC      120
165 TTTTATGGTG TATGCAATGT CATTCTATC ACTTGTGCTT ATACATCATT TTCAATAGTT      180
167 CCTCCCAGCA ATGGGACCAC TATGTGGAGA GCCACAATA CAGTCTTCAG TGCCATTTTG      240
169 GCTTTTTTAC TCGTAGATGA GAAAATGGCT TATGTTGACA TGGCTACAGT TGTTTGACAGC      300
171 ATCTTAGGTG TTTGTCTTGT CATGATCCCA AACATTGTTG ATGAAGACAA TTCTTTGTTA      360
173 AATGCCTGGA AAGAAGCCTT TGGGTACACC ATGACTGTGA TGGCTGGACT GACCACTGCT      420
175 CTCTCAATGA TAGTATACAG ATCCATCAAG GAGAAGATCA GCATGTGGAC TGCACTGTTT      480
177 ACTTTTGGTT GGACTGGGAC AATTTGGGGA ATATCTACTA TGTTTATTCT TCAAGAACCC      540
179 ATCATCCCAT TAGATGGAGA AACCTGGAGT TATCTCATTG CTATATGTGT CTGTTCTACT      600
181 GCAGCATTCT TAGGAGTTTA TTATGCCTTG GACAAATTCC ATCCAGCTTT GGTAGCACA      660
183 GTACAACATT TGGAGATTGT GGTAGCTATG GTCTTGACG TTCTCGTGCT GCACATATTT      720
185 CCTAGCATCT ATGATGTTTT TGGAGGGGTA ATCATTATGA TTAGTGTTTT TGTCCTTGCT      780
187 GGCTATAAAC TTTACTGGAG GAATTTAAGA AGGCAGGACT ACCAGGAAAT ATTAGACTCT      840
189 CCCATTAAAT GAATACCTGA TTATTATTGT CTCATTAATG TTCAGTTATT AATATGTATA      900
191 CTGCCATTTT AATGTTTACC TATGAATGTC TTTTGTGTTA TATAACTGAC AGAGTGCTAT      960
193 AAAATATATA ATATATACAA ATGCAGAAAA TTTATTCTAG TCTAATATAT TCAAAATACAA     1020
195 ATATTAAATA TATGAAATAC GTTAAAAAAA AAAAAAAAAA AAA                      1063

```

198 (2) INFORMATION FOR SEQ ID NO: 4:

200 (i) SEQUENCE CHARACTERISTICS:

201 (A) LENGTH: 216 amino acids

202 (B) TYPE: amino acid

203 (C) STRANDEDNESS:

204 (D) TOPOLOGY: linear

206 (ii) MOLECULE TYPE: protein

209 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```

211 Met Trp Arg Ala Thr Thr Thr Val Phe Ser Ala Ile Leu Ala Phe Leu
212 1          5          10          15
214 Leu Val Asp Glu Lys Met Ala Tyr Val Asp Met Ala Thr Val Val Cys
215          20          25          30
216 Ser Ile Leu Gly Val Cys Leu Val Met Ile Pro Asn Ile Val Asp Glu
217          35          40          45
219 Asp Asn Ser Leu Leu Asn Ala Trp Lys Glu Ala Phe Gly Tyr Thr Met
220          50          55          60
222 Thr Val Met Ala Gly Leu Thr Thr Ala Leu Ser Met Ile Val Tyr Arg
223          65          70          75          80
225 Ser Ile Lys Glu Lys Ile Ser Met Trp Thr Ala Leu Phe Thr Phe Gly
226          85          90          95
228 Trp Thr Gly Thr Ile Trp Gly Ile Ser Thr Met Phe Ile Leu Gln Glu
229          100         105         110
231 Pro Ile Ile Pro Leu Asp Gly Glu Thr Trp Ser Tyr Leu Ile Ala Ile
232          115         120         125
234 Cys Val Cys Ser Thr Ala Ala Phe Leu Gly Val Tyr Tyr Ala Leu Asp
235          130         135         140
237 Lys Phe His Pro Ala Leu Val Ser Thr Val Gln His Leu Glu Ile Val
238          145         150         155         160
240 Val Ala Met Val Leu Gln Leu Leu Val Leu His Ile Phe Pro Ser Ile
241          165         170         175
243 Tyr Asp Val Phe Gly Gly Val Ile Ile Met Ile Ser Val Phe Val Leu
244          180         185         190

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/745,763

DATE: 07/18/2001

TIME: 12:08:34

Input Set : A:\Cpg.pto

Output Set: N:\CRF3\07182001\I745763.raw

```

246   Ala Gly Tyr Lys Leu Tyr Trp Arg Asn Leu Arg Arg Gln Asp Tyr Gln
247           195                200                205
249   Glu Ile Leu Asp Ser Pro Ile Lys
250       210                215
254 (2) INFORMATION FOR SEQ ID NO: 5:
256   (i) SEQUENCE CHARACTERISTICS:
257       (A) LENGTH: 356 base pairs
258       (B) TYPE: nucleic acid
259       (C) STRANDEDNESS: double
260       (D) TOPOLOGY: linear
262   (ii) MOLECULE TYPE: cDNA
265   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
267 TGGCCAAAGA GGCCTAGCCG GGAGCGGGCG AGGCGGCGGC GGCAGCAGCG ATGGCAGGAA      60
269 TAGAGTTGGA GCGGTGCCAG CAGCAGGCCGA ACGAGGTGAC GGAAATTATG CGTAACAACT      120
271 TCGGCAAGGT CCTGGAGCGT GGTGTGAAGC TGGCCGAAC TGCAGCAGCGT TCAGACCAAC      180
273 TCCTGGATAT GAGCTCAACC TTCAACAAGA CTACACAGAA CCTGGCCCAG AAGAAGTGCT      240
275 GGGAGAACAT CCGTTACCGG ATCTGCGTGG GGCTGGTGGT GGTGTTGGTCTC CTGCTCATCA      300
277 TCCTGATTGT GCTGCTGGTC GTCTTTCTCC CTCAGAGCAG TGACAGCAGT AGTGCC      356
280 (2) INFORMATION FOR SEQ ID NO: 6:
282   (i) SEQUENCE CHARACTERISTICS:
283       (A) LENGTH: 102 amino acids
284       (B) TYPE: amino acid
285       (C) STRANDEDNESS:
286       (D) TOPOLOGY: linear
288   (ii) MOLECULE TYPE: protein
291   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
293   Met Ala Gly Ile Glu Leu Glu Arg Cys Gln Gln Gln Ala Asn Glu Val
294     1           5           10           15
296   Thr Glu Ile Met Arg Asn Asn Phe Gly Lys Val Leu Glu Arg Gly Val
297     20           25           30
299   Lys Leu Ala Glu Leu Gln Gln Arg Ser Asp Gln Leu Leu Asp Met Ser
300     35           40           45
302   Ser Thr Phe Asn Lys Thr Thr Gln Asn Leu Ala Gln Lys Lys Cys Trp
303     50           55           60
305   Glu Asn Ile Arg Tyr Arg Ile Cys Val Gly Leu Val Val Val Gly Val
306     65           70           75           80
308   Leu Leu Ile Ile Leu Ile Val Leu Leu Val Val Phe Leu Pro Gln Ser
309           85           90           95
311   Ser Asp Ser Ser Ser Ala
312           100
315 (2) INFORMATION FOR SEQ ID NO: 7:
317   (i) SEQUENCE CHARACTERISTICS:
318       (A) LENGTH: 92 base pairs
319       (B) TYPE: nucleic acid
320       (C) STRANDEDNESS: double
321       (D) TOPOLOGY: linear
323   (ii) MOLECULE TYPE: cDNA
325   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
327 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA      60

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/745,763

DATE: 07/18/2001

TIME: 12:08:34

Input Set : A:\Cpg.pto

Output Set: N:\CRF3\07182001\I745763.raw

```

329 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AA                                     92
332 (2) INFORMATION FOR SEQ ID NO: 8:
334     (i) SEQUENCE CHARACTERISTICS:
335         (A) LENGTH: 1131 base pairs
336         (B) TYPE: nucleic acid
337         (C) STRANDEDNESS: double
338         (D) TOPOLOGY: linear
340     (ii) MOLECULE TYPE: cDNA
343     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
345 GGGCCTCAAC TTTGGCGTCG TGAGATTCTT GTGAGGCGTC TGCCTGGAAG CCGGCAGCAA      60
347 TTTTGCTTCT TTAAAGAGAA AAAGAAGGCT AGGGACTCAG ATTCCTGGAT TCTGAGATCC      120
349 AGACCAGCTC CTCCAGACC TCTCCAGAAG AAGCCATGGG AACCCTCGT ATCCAGCATT      180
351 TGCTGATCCT CCTGGTCTTA GGAGCCTCCC TCCTGACCTC GGGCCTAGAG CTGTATTGTC      240
353 AAAAGGGTCT GTCCATGACT GTGGAAGCAG ATCCAGCCAA TATGTTAAC TGGACCACAG      300
355 AGGAAGTGGA GACTTGTGAC AAAGGGGCAC TTTGCCAGGA AACCATACTA ATAATTAAAG      360
357 CAGGGACTGA GACAGCCATT TTGGCCACGA AGGGCTGCAT CCCGGAAGGG GAGGAGGCCA      420
359 TAACAATTGT CCAGCACTCT TCACCTCCCG GCCTGATCGT GACCTCCTAC AGTAACTACT      480
361 GTGAGGATTC CTTCTGTAAT GACAAAGACA GCCTGTCTCA GTTTGGGAG TTCAGTGAGA      540
363 CCACAGCTTC CACTGTGTCA ACAACCCTCC ATTGTCCAAC CTGTGTGGCT TTGGGGACCT      600
365 GTTTCAGTGC TCCTTCTCTT CCCTGTCCCA ATGGTACAAC TCGATGCTAT CAAGGAAAAC      660
367 TTGAGATCAC TGGAGGTGGC ATTGATCGT CTGTGGAGGT CAAAGGCTGT ACAGCCATGA      720
369 TTGGCTGCAG GCTGATGTCT GGAATCTTAG CAGTAGGACC CATGTTGTG AGGGAAGCGT      780
371 GCCACATCA GCTGCTCACT CAACCTCGAA AGACTGAAAA TGGGGCCACC TGTCTTCCCA      840
373 TTCCTGTTTG GGGGTACAG CTACTGCTGC CATTGCTGCT GCCATCATTT ATTCACTTTT      900
375 CCTAAGAAGG CACTTCTGGG CCTGGGTCTG AGGACATCTT TTTTGACTGG GAGCCTTCTT      960
377 ACTGTTGAGG TTCAACAAGC TGAGGAGTAG ATGGGAATTT GAGGGAGAAT ACAGAGATAC     1020
379 TATGAACGTA TTTGACATTT TTAATACAAT TTCTGCTATA ATTTTGTAT GCAGTAGGCG     1080
381 TTAATAATAA ACATTCTGCT TGTGAAAAAA AAAAAAAAAA AAAAAAAAAA A      1131
385 (2) INFORMATION FOR SEQ ID NO: 9:
387     (i) SEQUENCE CHARACTERISTICS:
388         (A) LENGTH: 249 amino acids
389         (B) TYPE: amino acid
390         (C) STRANDEDNESS:
391         (D) TOPOLOGY: linear
393     (ii) MOLECULE TYPE: protein
395     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:
397 Met Gly Thr Pro Arg Ile Gln His Leu Leu Ile Leu Leu Val Leu Gly
398 1           5           10           15
400 Ala Ser Leu Leu Thr Ser Gly Leu Glu Leu Tyr Cys Gln Lys Gly Leu
401           20           25           30
403 Ser Met Thr Val Glu Ala Asp Pro Ala Asn Met Phe Asn Trp Thr Thr
404           35           40           45
406 Glu Glu Val Glu Thr Cys Asp Lys Gly Ala Leu Cys Gln Glu Thr Ile
407           50           55           60
409 Leu Ile Ile Lys Ala Gly Thr Glu Thr Ala Ile Leu Ala Thr Lys Gly
410           65           70           75           80
412 Cys Ile Pro Glu Gly Glu Glu Ala Ile Thr Ile Val Gln His Ser Ser
413           85           90           95
415 Pro Pro Gly Leu Ile Val Thr Ser Tyr Ser Asn Tyr Cys Glu Asp Ser

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/745,763

DATE: 07/18/2001

TIME: 12:08:35

Input Set : A:\Cpg.pto

Output Set: N:\CRF3\07182001\I745763.raw

L:35 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:36 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]

L:1073 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:17

L:1076 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:17

L:2370 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:42

L:2411 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:42

L:2701 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:48

L:2793 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:50

L:3168 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:63

L:4507 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:87

L:4524 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:87

L:4900 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:100

L:5605 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:108

L:5608 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:108

L:5632 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:108

L:5638 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:108

L:6050 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:115

L:6154 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:115

L:6157 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:115

L:7303 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:140

L:7312 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:140

L:10459 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:191

L:10540 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:191

L:10835 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:197

L:11742 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:207

L:11745 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:207

L:11754 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:207

L:11766 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:207

OIPE

RAW SEQUENCE LISTING

DATE: 06/28/2001

PATENT APPLICATION: US/09/745,763

TIME: 14:39:43

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\06282001\I745763.raw

SEQUENCE LISTING

Does Not Comply
Corrected Diskette Needed

pp. 1, 4

4 (1) GENERAL INFORMATION:

6 (i) APPLICANT: Jacobs, Kenneth

7 McCoy, John M.

8 LaVallie, Edward R.

9 Collins-Racie, Lisa A.

10 Evans, Cheryl

11 Merberg, David

12 Treacy, Maurice

13 Spaulding, Vikki

15 (ii) TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES

16 ENCODING THEM

18 (iii) NUMBER OF SEQUENCES: 219

20 (iv) CORRESPONDENCE ADDRESS:

21 (A) ADDRESSEE: Genetics Institute, Inc.

22 (B) STREET: 87 CambridgePark Drive

23 (C) CITY: Cambridge

24 (D) STATE: MA

25 (E) COUNTRY: U.S.A.

26 (F) ZIP: 02140

28 (v) COMPUTER READABLE FORM:

29 (A) MEDIUM TYPE: Floppy disk

30 (B) COMPUTER: IBM PC compatible

31 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

32 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

34 (vi) CURRENT APPLICATION DATA:

C--> 35 (A) APPLICATION NUMBER: US/09/745,763

C--> 36 (B) FILING DATE: 18-Jun-2000

37 (C) CLASSIFICATION:

39 (viii) ATTORNEY/AGENT INFORMATION:

40 (A) NAME: Sprunger, Suzanne A.

41 (B) REGISTRATION NUMBER: 41,323

43 (ix) TELECOMMUNICATION INFORMATION:

44 (A) TELEPHONE: (617) 498-8284

45 (B) TELEFAX: (617) 876-5851

ERRORED SEQUENCES

1667 (2) INFORMATION FOR SEQ ID NO: 33:

1669 (i) SEQUENCE CHARACTERISTICS:

1670 (A) LENGTH: 2199 base pairs

1671 (B) TYPE: nucleic acid

1672 (C) STRANDEDNESS: double

1673 (D) TOPOLOGY: linear

1675 (ii) MOLECULE TYPE: cDNA

E--> 1677 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:

Format error at sequence
33.

RAW SEQUENCE LISTING

DATE: 06/28/2001

PATENT APPLICATION: US/09/745,763

TIME: 14:39:43

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\06282001\I745763.raw

```

1679 AGTTGGCAGG TGGAGAGGCA GGTGAGGAGG GAAAGTCGGG GGAGGACGCG GAAGAGGAGC      60
1681 TGTGGGAAGG GGGAGGAGGG AGGGAGGAAA AGAGGAGGAG GCGGAGGAGA ACTGAGCAGA      120
1683 GCAGAGCATC GAGCCAAAGG GGAGATGAGT TTGTCTGTCC TCTGCTGAGG CTACGGCCGG      180
1685 GCCTAGGGAA CTGGGAGCTT GGGTGGGAGC GACACCCGTG GAAGTGGGAG GAGGTGGCGC      240
1687 CGGGACTTTA ACCCCTTGTG GGCTCTGCGG CAGGGGATTT AACCCCTTGT GGATCTGGCC      300
1689 CCTCGGAGGC AGCGTCATCG GTAGTTTTAA CCCCTTCGGG GCTGGGTTTC ACGCACTGGA      360
1691 CTTACCTCA TCACCTTGCT CACCAACTCC TTTATTGGGG TGCTCCGCTT GGAGGTTTGA      420
1693 GGCCACCTC CGCCATTAC GTACTGTTCC TGCCGCTGCA CCCCTTGGA CCGCTAGCT      480
1695 GGCCGCACTG TGGGCGCTTA ACCCTTTACT GACTTGAGCT CCCCAGATTG CAGTTGGAGT      540
1697 TTGCTGATAG AAGGACTAGC TAAAGGCGTC ACTGCAGGAA TTACAAACTG AAGAGGACTC      600
1699 TGTGGACTG TTTTTTTTTT CTTTTCTTT TTTTAAAGAA AAACCCATTT TTTTCTTAA      660
1701 GGACTTACTA GCCAAAATTT CTTAACTTC GAGGACTCTA CTAGCCATGG CCGAGCCATT      720
1703 CTGTGAGAA TATCAACACC AGCCTCAAAC TAGCAACTGT ACAGGTGCTG CTGCTGTCCA      780
1705 GGAAGAGCTG AACCTGAGC GCCCCCCAGG CGCGGAGGAG CGGGTGCCCG AGGAGGACAG      840
1707 TAGGTGGCAA TCGAGAGCGT TCCCCAGTT GGGTGGCCGT CCGGGGCCGG AGGGGAAGG      900
1709 GAGCCTGGAA TCCCAACCAC CTCCCTTGCA GACCCAGGCC TGTCCAGAAT CTAGCTGCCT      960
1711 GAGAGAGGGC GAGAAGGGCC AGAATGGGGA CGACTCGTCC GCTGGCGGCG ACTTCCCGCC      1020
1713 GCCGGCAGAA GTGGAACCGA CGCCCGAGGC CGAGCTGCTC GCCCAGCCTT GTCATGACTC      1080
1715 CGAGGCCAGT AAGTTGGGGG CTCCTGCCGC AGGGGGCGAA GAGGAGTGGG GACAGCAGCA      1140
1717 GAGACAGCTG GGAAGAAAA AACATAGGAG ACGCCCGTCC AAGAAGAAGC GGCATTGGAA      1200
1719 ACCGTACTAC AAGCTGACCT GGAAGAGAA GAAAAAGTTC GACGAGAAAC AGAGCCTTCG      1260
1721 AGCTTCAAGG ATCCGAGCCG AGATGTTGCG CAAGGGCCAG CCGGTCGCGC CCTATAACAC      1320
1723 CACGCAGTTC CTCATGGATG ATCACGACCA GGAGGAGCCG GATCTCAAAA CCGGCCTGTA      1380
1725 CTCCAAGCGG GCCGCCGCCA AATCCGACGA CACCAGCGAT GACGACTTCA TGAAGAAGG      1440
1727 GGGTGAGGAG GATGGGGGCA GCGATGGGAT GGGAGGGGAC GGCAGCGAGT TTCTGCAGCG      1500
1729 GGACTTCTCG GAGACGTACG AGCGGTACCA CACGGAGAGC CTGCAGAACA TGAGCAAGCA      1560
1731 GGAGCTCATC AAGGAGTACC TGGAAGTGA GAAGTGCCTC TCGCGCATGG AGGACGAGAA      1620
1733 CAACCGGCTG CGGCTGGAGA GCAAGCGGCT GGGTGGCGAG GACGCGCGTG TCGGGGAGCT      1680
1735 GGAGCTGGAG CTGGACCGGC TGCGCGCCGA GAACCTCCAG CTGCTGACCG AGAACGAACT      1740
1737 GCACCGGAGC CAGGAGCGAG CGCCGCTTTC CAAGTTTGA GACTAGACTG AAACCTTTT      1800
1739 GGGGGAGGGG GCAAAGGGGA CTTTTTACAG TGATGGAATG TAACATTATA TACATGTGTA      1860
1741 TATAAGACAG TGGACCTTTT TATGACACAT AATCAGAAGA GAAATCCCCC TGGCTTTGGT      1920
1743 TGGTTTCGTA AATTTAGCTA TATGTAGCTT GCGTGCTTTC TCCTGTTCTT TTAATTATGT      1980
1745 GAAACTGAAG AGTTGCTTTT CTTGTTTTCC TTTTLAGAAG TTTTTTTCCT TAATGTGAAA      2040
1747 GTAATTTGAC CAAGTTATAA TGCATTTTTG TTTTAAACAA ATCCCCTCCT TAAACGGAGC      2100
1749 TATAAGGTGG CCAAATCTGA GAACAATTAA ATTCATTTTA GTTATAATAA ATTTAATATT      2160
1751 TGTAAATGTA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA      2199

```

12067 (2) INFORMATION FOR SEQ ID NO: 219:

12069 (i) SEQUENCE CHARACTERISTICS:

12070 (A) LENGTH: 542 amino acids

12071 (B) TYPE: amino acid

12072 (C) STRANDEDNESS:

12073 (D) TOPOLOGY: linear

12075 (ii) MOLECULE TYPE: protein

12077 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 219:

12079 Cys Gly His His Glu Leu Asn Asn Leu Thr Gln Val Gln Gln

12080 1 5 10 15

12082 Arg Asn Leu Ile Thr Asn Leu Gln Arg Ser Val Asp Asp Thr Ser Gln

12083 20 25 30

RAW SEQUENCE LISTING

DATE: 06/28/2001

PATENT APPLICATION: US/09/745,763

TIME: 14:39:45

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\06282001\I745763.raw

12085	Ala	Ile	Gln	Arg	Ile	Lys	Asn	Asp	Phe	Gln	Asn	Leu	Gln	Gln	Val	Phe
12086			35					40					45			
12088	Leu	Gln	Ala	Lys	Lys	Asp	Thr	Asp	Trp	Leu	Lys	Glu	Lys	Val	Gln	Ser
12089			50				55					60				
12091	Leu	Gln	Thr	Leu	Ala	Ala	Asn	Asn	Ser	Ala	Leu	Ala	Lys	Ala	Asn	Asn
12092			65				70					75			80	
12094	Asp	Thr	Leu	Glu	Asp	Met	Asn	Ser	Gln	Leu	Asn	Ser	Phe	Thr	Gly	Gln
12095					85					90					95	
12097	Met	Glu	Asn	Ile	Thr	Thr	Ile	Ser	Gln	Ala	Asn	Glu	Gln	Asn	Leu	Lys
12098				100					105					110		
12100	Asp	Leu	Gln	Asp	Leu	His	Lys	Asp	Ala	Glu	Asn	Arg	Thr	Ala	Ile	Lys
12101			115					120					125			
12103	Phe	Asn	Gln	Leu	Glu	Glu	Arg	Phe	Gln	Leu	Phe	Glu	Thr	Asp	Ile	Val
12104			130				135					140				
12106	Asn	Ile	Ile	Ser	Asn	Ile	Ser	Tyr	Thr	Ala	His	Leu	Arg	Thr	Leu	
12107					145		150				155				160	
12109	Thr	Ser	Asn	Leu	Asn	Glu	Val	Arg	Thr	Thr	Cys	Thr	Asp	Thr	Leu	Thr
12110					165					170					175	
12112	Lys	His	Thr	Asp	Asp	Leu	Thr	Ser	Leu	Asn	Asn	Thr	Leu	Ala	Asn	Ile
12113				180					185					190		
12115	Arg	Leu	Asp	Ser	Val	Ser	Leu	Arg	Met	Gln	Gln	Asp	Leu	Met	Arg	Ser
12116			195					200					205			
12118	Arg	Leu	Asp	Thr	Glu	Val	Ala	Asn	Leu	Ser	Val	Ile	Met	Glu	Glu	Met
12119			210				215					220				
12121	Lys	Leu	Val	Asp	Ser	Lys	His	Gly	Gln	Leu	Ile	Lys	Asn	Phe	Thr	Ile
12122			225				230					235			240	
12124	Leu	Gln	Gly	Pro	Pro	Gly	Pro	Arg	Gly	Pro	Arg	Gly	Asp	Arg	Gly	Ser
12125				245						250					255	
12127	Gln	Gly	Pro	Pro	Gly	Pro	Thr	Gly	Asn	Lys	Gly	Gln	Lys	Gly	Glu	Lys
12128				260					265					270		
12130	Gly	Glu	Pro	Gly	Pro	Pro	Gly	Pro	Ala	Gly	Glu	Arg	Gly	Pro	Ile	Gly
12131			275					280					285			
12133	Pro	Ala	Gly	Pro	Pro	Gly	Glu	Arg	Gly	Gly	Lys	Gly	Ser	Lys	Gly	Ser
12134			290				295					300				
12136	Gln	Gly	Pro	Lys	Gly	Ser	Arg	Gly	Ser	Pro	Gly	Lys	Pro	Gly	Pro	Gln
12137			305				310				315				320	
12139	Gly	Pro	Ser	Gly	Asp	Pro	Gly	Pro	Pro	Gly	Pro	Pro	Gly	Lys	Glu	Gly
12140				325						330					335	
12142	Leu	Pro	Gly	Pro	Gln	Gly	Pro	Pro	Gly	Phe	Gln	Gly	Leu	Gln	Gly	Thr
12143				340					345					350		
12145	Val	Gly	Glu	Pro	Gly	Val	Pro	Gly	Pro	Arg	Gly	Leu	Pro	Gly	Leu	Pro
12146			355					360					365			
12148	Gly	Val	Pro	Gly	Met	Pro	Gly	Pro	Lys	Gly	Pro	Pro	Gly	Pro	Pro	Gly
12149			370				375					380				
12151	Pro	Ser	Gly	Ala	Val	Val	Pro	Leu	Ala	Leu	Gln	Asn	Glu	Pro	Thr	Pro
12152					385		390				395				400	
12154	Ala	Pro	Glu	Asp	Asn	Ser	Cys	Pro	Pro	His	Trp	Lys	Asn	Phe	Thr	Asp
12155				405						410					415	
12157	Lys	Cys	Tyr	Tyr	Phe	Ser	Val	Glu	Lys	Glu	Ile	Phe	Glu	Asp	Ala	Lys

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/745,763

DATE: 06/28/2001

TIME: 14:39:45

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\06282001\I745763.raw

```

12158          420          425          430
12160      Leu Phe Cys Glu Asp Lys Ser Ser His Leu Val Phe Ile Asn Thr Arg
12161          435          440          445
12163      Glu Glu Gln Gln Trp Ile Lys Lys Gln Met Val Gly Arg Glu Ser His
12164          450          455          460
12166      Trp Ile Gly Leu Thr Asp Ser Glu Arg Glu Asn Glu Trp Lys Trp Leu
12167      465          470          475          480
12169      Asp Gly Thr Ser Pro Asp Tyr Lys Asn Trp Lys Ala Gly Gln Pro Asp
12170          485          490          495
12172      Asn Trp Gly His Gly His Gly Pro Gly Glu Asp Cys Ala Gly Leu Ile
12173          500          505          510
12175      Tyr Ala Gly Gln Trp Asn Asp Phe Gln Cys Glu Asp Val Asn Asn Phe
12176          515          520          525
12178      Ile Cys Glu Lys Asp Arg Glu Thr Val Leu Ser Ser Ala Leu
12179          530          535          540
E--> 12180  - 165 -

```

→ Delete extraneous numeral at
the end of the file.

VERIFICATION SUMMARY

DATE: 06/28/2001

PATENT APPLICATION: US/09/745,763

TIME: 14:39:46

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\06282001\I745763.raw

L:35 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:36 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:1073 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:17
L:1076 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:17
L:1677 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]
L:0 M:200 E: Mandatory Header Field missing, SeqNo=33, SEQUENCE DESCRIPTION: SEQ ID NO: of (2)
L:1677 M:202 E: (16) Value must be an Integer, Data=[]
L:2370 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:42
L:2411 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:42
L:2701 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:48
L:2793 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:50
L:3168 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:63
L:4507 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:87
L:4524 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:87
L:4900 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:100
L:5605 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:108
L:5608 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:108
L:5632 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:108
L:5638 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:108
L:6050 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:115
L:6154 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:115
L:6157 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:115
L:7303 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:140
L:7312 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:140
L:10459 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:191
L:10540 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:191
L:10835 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:197
L:11742 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:207
L:11745 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:207
L:11754 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:207
L:11766 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:207
L:12180 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:219